

WHAT IS CLAIMED IS:

1. A speech synthesizer having an output stage for converting a phonetic description to an acoustic output, the output stage including a database of recorded utterance segments, in which the output stage:

- a. converts the phonetic description to a plurality of time-varying parameters;
- b. interprets the parameters as a key for accessing the database to identify an utterance segment in the database, and
- c. outputs the identified utterance segment;

in which the output stage further comprises an output waveform synthesizer that can generate an output signal from the parameters, whereby, in the event that the parameters describe an utterance segment for which there is no corresponding recording in the database, the parameters are passed to the output waveform synthesizer to generate an output signal.

2. A speech synthesizer according to claim 1 in which the output waveform synthesizer is essentially the same as the synthesizer used in a conventional parametric synthesizer.
3. A speech synthesizer according to claim 1 in which the database is populated to achieve a compromise between quality and memory requirement most appropriate to a specific application.
4. A speech synthesizer according to claim 3 in which the database is populated with segments that are most likely to be required to generate a range of output corresponding to the application of the synthesizer.
5. A speech synthesizer according to claim 4 in which the database is populated with utterance segments derived from speech by a particular individual speaker.

6. A speech synthesizer according to claim 4 in which the database is populated with utterance segments derived from speech by speakers of a particular gender.
7. A speech synthesizer according to claim 4 in which the database is populated with utterance segments derived from speech by speakers having a particular accent.
8. A speech synthesizer according to claim 1 in which the database is an indexed database.
9. A speech synthesizer according to claim 8 in which the index values for accessing the database are the values of the time-varying parameters.
10. A speech synthesizer according to claim 1 in which the segments within the database are coded.
11. A speech synthesizer according to claim 10 in which the segments within the database are coded using linear predictive coding, GSM coding or other coding schemes.
12. A speech synthesizer according to claim 1 in which the parameters are generated in regular periodic frames.
13. A speech synthesizer according to claim 12 in which the frames have a period of 2 to 30 ms.
14. A speech synthesizer according to claim 13 in which the period is approximately 10 ms.
15. A speech synthesizer according to claim 13 in which at each frame, an output waveform is generated these being reproduced in succession to create an impression of a continuous output.
16. A speech synthesizer according to claim 1 in which the parameters correspond to speech formants.

17. A method of synthesizing speech comprising:

- a. generating from a phonetic description a plurality of time-varying parameters that describe an output waveform;
- b. interpreting the parameters to identify an utterance segment within a database of such segments that corresponds to the audio output defined by the parameters and retrieving the segment to create an output waveform; and
- c. outputting the output waveform;

in which, if no utterance segment is identified in the database in step *b*, as corresponding to the parameters, an output waveform for output in step *c* is generated by synthesis.

- 18. A method of synthesizing speech according to claim 17 in which steps *a* to *c* are repeated in quick succession to create an impression of a continuous output.
- 19. A method of synthesizing speech according to claim 17 in which the parameters are generated in discrete frames, and steps *a* to *c* are performed once for each frame.
- 20. A method of synthesizing speech according to claim 17 in which the frames are generated with a regular periodicity.
- 21. A method of synthesizing speech according to claim 20 in which the frames are generated with a period of several ms (e.g. 10 ms or thereabouts).
- 22. A method of synthesizing speech according to claim 17 in which the parameters within the frames correspond to speech formants.
- 23. A method of synthesizing speech according to claims 17 in which the output segments for any one frame are selected as a function of the parameters of several frames.